



# Extraction of Temporal Network Structures from Graph-based Signals

Ronan Hamon, Pierre Borgnat, Patrick Flandrin, Céline Robardet

## ► To cite this version:

Ronan Hamon, Pierre Borgnat, Patrick Flandrin, Céline Robardet. Extraction of Temporal Network Structures from Graph-based Signals. IEEE Transactions on Signal and Information Processing over Networks, 2016, 2 (2), pp.215-226. 10.1109/TSIPN.2016.2530562 . hal-01330184

**HAL Id: hal-01330184**

**<https://hal.science/hal-01330184>**

Submitted on 10 Jun 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Extraction of Temporal Network Structures from Graph-based Signals

Ronan Hamon, Pierre Borgnat, *Member IEEE*, Patrick Flandrin, *Fellow IEEE*, and Céline Robardet

**Abstract**—A new framework to track the structure of temporal networks with a signal processing approach is introduced. The method is based on the duality between static networks and signals, obtained using a multidimensional scaling technique, that makes possible the study of the network structure from frequency patterns of the corresponding signals. In this paper, we propose an approach to identify structures in temporal networks by extracting the most significant frequency patterns and their activation coefficients over time, using nonnegative matrix factorization of the temporal spectra. The framework, inspired by audio decomposition, allows transforming back these frequency patterns into networks, to highlight the evolution of the underlying structure of the network over time. The effectiveness of the method is first evidenced on a synthetic example, prior being used to study a temporal network of face-to-face contacts. The extracted sub-networks highlight significant structures decomposed on time intervals that validates the relevance of the approach on real-world data.

**Index Terms**—temporal networks, multidimensional scaling, nonnegative matrix factorization, network structures, decomposition

## I. INTRODUCTION

Many complex systems, whether physical, biological or social, can be naturally represented as networks, i.e., a set of relationships between entities. Network science [1] has been widely developed to study such objects by providing powerful tools, for instance the detection of communities [2], in order to understand the underlying properties of these systems. These works benefit from the natural representation of networks by graphs, enabling the use of the comprehensive mathematical understanding of graph theory. Recently, connections between signal processing and network theory have emerged: The field of signal processing over networks [3] [4] has been introduced with the objective of transposing concepts developed in classical signal processing, such as Fourier transform or wavelets, in the graph domain. These works have led to significant results, among them filtering of signals defined over a network [3] or multiscale community mining using graph wavelets [5]. Connections have also been considered by defining a duality between graphs and signals: methods have been developed to transform graphs into signals and conversely, in order to take advantage of both signal processing and graph theory in the analysis. In this matter, mapping a graph into time series has been performed using random

walks [6] [7] [8] or deterministic methods based on classical multidimensional scaling [9]. This latter approach has been the topic of several extensions in [10], in order to build a comprehensive framework to link frequency patterns of the so-obtained signals with network structures.

Studies mainly focused on the analysis of static networks, potentially aggregated over time, but without time evolution. However, the considered systems are most of the time not frozen: vertices and edges appear and disappear over the course of time. Aggregating the networks over a time interval gives insight of the underlying mechanisms, but often does not provide the actual dynamic sequence of appearance and disappearance of edges: two edges, active one after the other, will be considered simultaneous in the temporally-aggregated network, leading to significant change of the network structures. Given the importance of knowing such dynamics, for instance in topics such as epidemic spread or communication networks, and thanks to the recent availability of many data sets, a temporal network theory has recently appeared [11] [12], developing tools to better understand the underlying mechanisms of systems. Several studies proposed an extension of the methods developed for static networks to the temporal case: we can cite for instance works on network modeling [13] [14], [15] [16], detection of communities [17] [18] [19] [20] [21] [22], detection of temporal motifs [23], visualization [24], or more generally data mining of time-evolving sequences [25].

We propose in this article a new approach based on the graph-based signals introduced in [10] to study temporal networks. It makes possible to visually track temporal networks by following the frequency patterns associated to specific structures. These frequency patterns are then studied by automatically extracting and tracking the significant frequency patterns over time using nonnegative matrix factorization (NMF) [26]. These patterns can be then transformed back into networks to reveal the most significant structures of the temporal network over time, the objective being here to follow the global evolution of the system.

Preliminary versions of this work have been presented in previous works: the principle of the extension to the temporal case has been introduced in [27], as well as the visual tracking of frequency patterns representing structures. In [28] and [29], this approach has been used to make a study of a bike sharing system in Lyon, which is not mentioned in this article. Finally, the idea of the decomposition using NMF has been suggested in [30]. This paper extends however those previous works by detailing a comprehensive framework and setting out consistent arguments for the validity of the method in the

This work was supported by the programs ARC 5 and ARC 6 of the région Rhône-Alpes, the project Vél'Innov ANR-12-SOIN-0001-02.

R. Hamon, P. Borgnat, and P. Flandrin are with Univ Lyon, Ens de Lyon, Univ Claude Bernard, CNRS, Laboratoire de Physique, F-69342 Lyon, France.

C. Robardet is with Univ Lyon, INSA Lyon, Univ Claude Bernard, CNRS, LIRIS, F-69621 Villeurbanne, France.

context of analysis of temporal networks.

The paper is organized as follows. Section II briefly summarizes the duality between static networks and signals that has been defined in [9] and further studied in [10]. Section III is first concerned with the extension of the (static) duality to temporal networks. Then the spectral analysis of the obtained signals is introduced, enabling the use of different methods to visually track the structure of temporal networks. Finally the NMF-based method to extract significant frequency patterns over time is described, as well as how these frequency patterns can be transformed back into network structures. Section IV introduces a synthetic temporal network built as a succession of four different network structures which mix communities, cycles and randomness. It is shown that the method proposed in Section III permits to go beyond a mere visual tracking of the structure of the temporal network, revealing the different types of structures and evidencing their temporal activities in consistence with the construction of the synthetic temporal networks. A short study is also performed to assess the sensitivity of the temporal patterns according to the level of noise that may affect the original graph data. Finally, Section V provides an illustration on a real-world temporal network. The same methodology is applied to the well-documented case of a real-world social network that consists of face-to-face contacts in a primary school.

**Notations:** Matrices are denoted by boldface capital letters, their columns by boldface lowercase letters, and their elements by lowercase letters: for  $\mathbf{M} \in \mathbb{R}^{A \times B}$ ,  $\mathbf{M} = [\mathbf{m}_1, \dots, \mathbf{m}_B] = (m_{ab})_{a \in \{1, \dots, A\}, b \in \{1, \dots, B\}}$ . Tensors are represented by calligraphic capital letters: for  $\mathcal{M} \in \mathbb{R}^{A \times B \times T}$ ,  $\mathcal{M} = [\mathbf{M}^{(t)}]_{t=1, \dots, T}$ . Operators are represented by calligraphic boldface capital letters:  $\mathcal{F}$ .

## II. BACKGROUND ON GRAPH-BASED SIGNALS FOR STATIC NETWORKS

### A. Transformation from static networks to signals

Duality between networks and signals has been introduced to analyze networks structures using signal processing tools. Shimada et al. [9] proposed a method to transform static networks into a collection of signals using multidimensional scaling. In [10], this method has been extended to obtain a comprehensive framework to transform graphs into a collection of signals. As this method is central in the subsequent work on temporal networks, the main aspects of the method are recalled in this section. We consider in the following networks that are described by unweighted and undirected graphs with  $N$  vertices.

Classical MultiDimensional Scaling (CMDS) [31] is a statistical technique used to find the coordinates of points in a Euclidean space, whose only the relations of distance are known. This is done through a diagonalization of the distance matrix, previously double-centered: the eigenvectors are the corresponding signals, weighted by the square root of the corresponding eigenvalues. These signals are ordered by decreasing order of eigenvalue: the first components are then the most representative of the data structure.

In [9], CMDS is used to transform a graph into signals by projecting the  $N$  vertices of the graph in a Euclidean

space, such that distances between these points correspond to relations in the graph. Describing the graph by its adjacency matrix  $\mathbf{A} \in \mathbb{R}^{N \times N}$ , whose elements  $a_{ij}$  are equal to 1 if vertices  $i$  and  $j$  are linked, and 0 otherwise, the distance between vertices of a graph, denoted  $\Delta = (\delta_{ij})_{i,j=1, \dots, N}$ , is defined for two vertices  $i, j$  of the graph by:

$$\delta_{ij} = \begin{cases} 0 & \text{if } i = j \\ 1 & \text{if } a_{ij} = 1 \text{ and } i \neq j \\ w > 1 & \text{if } a_{ij} = 0 \text{ and } i \neq j \end{cases} \quad (1)$$

As discussed in [10], we choose  $w = 1 + \frac{1}{N}$ . This definition focuses on the presence, denoted by a distance equal to 1, or the absence, denoted by a distance equal to  $w$ , of an edge between two vertices. Hence, the distance of two vertices in the graph, often defined as the length of the shortest path between the two vertices, has no direct influence on the matrix  $\Delta$ : two unlinked vertices have a distance equal to  $w$ , whether they are close or not in the graph.

Applying CMDS on the distance matrix  $\Delta$  leads to a collection of points, corresponding to the vertices, in a Euclidean space  $\mathbb{R}^{N-1}$ . The obtained collection of signals is denoted by  $\mathbf{X} \in \mathbb{R}^{N \times C}$ , where  $C$  is the total number of components, and thus is equal to  $N - 1$  for the full representation. The columns  $\mathbf{x}_c$  represent the  $c$ -th signal, with  $c = 1, \dots, C$ , and are indexed by the vertices. The transformation is denoted by  $\mathcal{T}$ : for an adjacency matrix  $\mathbf{A}$ , we have  $\mathbf{X} = \mathcal{T}[\mathbf{A}]$ . Connections can be done between the obtained signals and the eigenvectors obtained by the diagonalization of the Laplacian matrix of the graph [10].

As the signals are indexed by the vertices, the order in which we consider them in the transformation is essential to study some aspects of the signals, especially when using spectral analysis of the signals. Ordering randomly the vertices does not change the value assigned to each vertex, but would lead to abrupt variations in the representation of signals: Specific frequency properties, clearly observable in signals, will no longer be visible. Unfortunately, the suitable ordering is usually not available, especially when dealing with real-world networks. To address this issue, we proposed in [32] to find a vertex ordering that reflects the topology of the underlying graph, based on the following assumption: if two vertices are close in the graph (by considering for instance the length of the shortest path between them), they have to be also close in the ordering. Details of the algorithm and results about the consistency between the obtained vertex ordering and the topology of graphs are entirely covered in [32].

### B. Inverse Transformation

Transforming back signals to a graph has to take into account the nature of the signals, as they are a representation of a specific network. The inverse transformation must hence preserve the original topology of the underlying graph. By construction of the collection of signals  $\mathbf{X}$ , the perfect retrieval of the underlying graph is easily reachable, by considering the distances between each pair of point: As built using CMDS, these distances represent the distance matrix  $\Delta$ , and the adjacency matrix of the graph is directly obtained.

However, when  $\mathbf{X}$  is degraded or modified, e.g., by retaining only a small number of components or by processing the signals, the distances are no longer directly the ones computed between vertices, even if they stay in the neighborhood of these distances. We proposed in [10] to take into account the energy of components to improve the reconstruction, as well as prior information about the original graph. If the distance between two vertices  $i$  and  $j$  in a high-energy component, i.e., a component associated to a large eigenvalue, is high, it means that the two vertices are likely to be distant in the graph. Conversely, if the distance is low, then the two vertices are likely to be connected in the graph.

Let  $\tilde{\mathbf{X}}$  be a degraded collection of signals. The energies are normalized according to the energies of the original components, by multiplying the components  $\tilde{x}_c$  by the normalization factor  $N_c$ :

$$N_c = \sqrt{\frac{\sum_{n=1}^N x_{nc}^2}{\sum_{n=1}^N \tilde{x}_{nc}^2}} \quad (2)$$

Then the distances are computed by using the energies as follows:

$$d(\mathbf{X})_{ij} = \sqrt{\sum_{c=1}^C (u_c)^\alpha (x_{ic} - x_{jc})^2} \quad (3)$$

with  $\alpha \geq 0$ , where  $u_c$  is the energy of component  $c$ , computed as  $u_c = \sum_{i=1}^n x_{ic}^2$ , and normalized such that  $\|\mathbf{u}\|_\alpha = \sum_{c=1}^C (u_c)^\alpha = C$ , with  $\mathbf{u}$  the vector of energies for all components. The parameter  $\alpha$  controls the importance of the weighting: if  $\alpha$  is high, the high-energy components have a higher importance in the computation of distances compared to the low-energy components. Conversely, if  $\alpha$  is small, the importance of high-energy components is diminished. In particular,  $\alpha = 0$  gives the standard reconstruction. The distances are then thresholded, by retaining the edges corresponding to the smallest distances. The threshold is chosen in order to recover the same amount of edges than in the original network.

It is worth noting that it is not strictly speaking an inverse transformation: if we start from a graph  $G$  and apply the method of transforming it into a collection of signals  $\mathbf{X} = \mathcal{T}[G]$ , this transformation is one-to-one, as is the following Fourier transformation that ends up with spectral features. Provided that the so-obtained signals (or spectra) are unaltered, the way back is guaranteed and the characterization can be considered as unique:  $\mathcal{T}^{-1}[\mathcal{T}[G]] = G$ . The situation however turns out to be different whenever the signals (or spectra) are altered prior reconstruction. In this case, applying the same inverse transformation  $\mathcal{T}^{-1}$  as before (i.e., just computing the distance matrix between components) on some altered signal  $\tilde{\mathbf{X}}$  does not result, in the general case, in a matrix which would be bimodal in its non-zero values and, hence, transformable into an admissible adjacency matrix. Indeed, this situation now requires to modify the reconstruction process, e.g., as proposed in [10], in which the modified inverse transform  $\tilde{\mathcal{T}}^{-1}$  results from the computing of a weighted distance matrix followed by a suitable thresholding, ending up with a graph  $\tilde{G} = \tilde{\mathcal{T}}^{-1}[\tilde{\mathbf{X}}]$  which can be an approximation of  $G$ , but whose transform

cannot be guaranteed to exactly coincide with the collection of signals we started with (in other words,  $\mathcal{T}[\tilde{\mathcal{T}}^{-1}[\tilde{\mathbf{X}}]] \neq \tilde{\mathbf{X}}$ ).

### III. EXTRACTION OF NETWORK STRUCTURES IN TEMPORAL NETWORKS

#### A. Transformation of temporal networks into signals and back

The description of temporal networks considered in this work consists in a discrete-time sequence of graph snapshots. The collection of all snapshots at the different times  $t = 0, \dots, T-1$ , with  $T$  the total number of time steps, can be represented by a graph adjacency tensor denoted  $\mathcal{A} \in \mathbb{R}^{N \times N \times T}$  such that  $\mathcal{A} = [\mathbf{A}^{(t)}]_{t=1, \dots, T}$ . We study here temporal networks where the edges are changing over time, keeping the same given set of vertices (possibly isolated).

The extension of the method described in Section II is directly achieved by applying at each time step the transformation on the corresponding static representation of the temporal network. We denote  $\mathcal{X} \in \mathbb{R}^{N \times C \times T}$ , the collection of signals obtained from  $\mathcal{A}$ , such that  $\mathcal{X} = [\mathbf{X}^{(t)}]_{t=1, \dots, T}$ . For each time step  $t$ , we have

$$\mathbf{X}^{(t)} = \mathcal{T}[\mathbf{A}^{(t)}] \quad (4)$$

As the number of vertices in the graph does not evolve, the number of components  $C$ , and then the number of frequencies  $F$ , is constant over time. In the case where the set of vertices evolves, the tensor is built by fixing the number of components as the maximal number of components over time, and by zero-padding the missing components. To index As for the indexation of signals, the algorithm relabeling vertices according to the structure of the network is performed at each time step, leading to different labelings of the vertices over time.<sup>1</sup>

Conversely, the inverse transformation is performed likewise, by applying the static inverse transformation on the collection of signals at time  $t$ :

$$\mathbf{A}^{(t)} = \mathcal{T}^{-1}[\mathbf{X}^{(t)}] \quad (5)$$

where  $\mathbf{A}^{(t)}$  represents the adjacency matrix of the reconstructed temporal network at time  $t$ .

#### B. Spectral analysis of signals

Spectral analysis is performed using standard signal processing methods: Let a collection  $\mathbf{X}$  of  $C$  signals indexed by  $N$  vertices. The spectra  $\mathbf{S} \in \mathbb{C}^{C \times F}$  give the complex Fourier coefficients, whose elements are obtained by applying the Fourier transform on each of the  $C$  components of  $\mathbf{X}$ :

$$\mathbf{s}_c = \mathcal{F}[\mathbf{x}_c] \quad (6)$$

estimated, for positive frequencies, on  $F = \frac{N}{2} + 1$  bins,  $\mathcal{F}$  being the Fourier transform and  $c = 1, \dots, C$ .

From the spectrum  $\mathbf{S}$ , the following features are obtained for each frequency of each component:

<sup>1</sup>The objective of this work here is to track how the structure of the temporal network evolves, regardless the labels of the vertices. The modification of the labeling indicates how the vertex evolves over time in the global structure of the network. This aspect is not discussed in the following, the labeling over time is kept only for purposes of reconstruction.



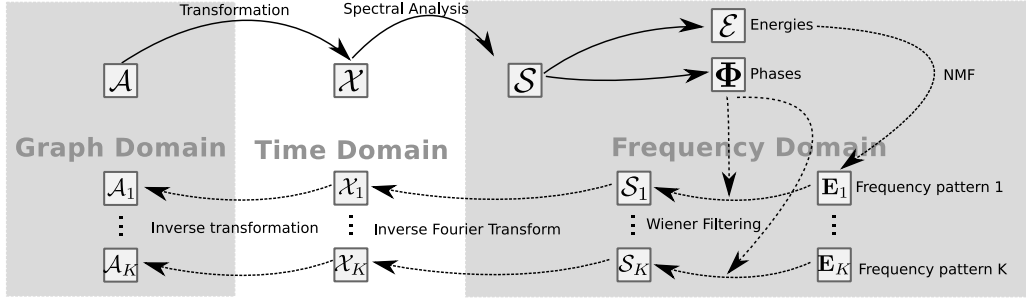


Fig. 1. This diagram describes the framework implemented to extract network structures in temporal networks using graph-based signals. The boxes describe the objects, while arrows represent the operations. The procedure is as follows: From a temporal network described by the adjacency tensor  $\mathcal{A}$ , a collection of signals  $\mathcal{X}$  is obtained by transforming each static graph into a collection of signals using the transformation described in [10]. A spectral analysis of the signals is then performed, and a tensor of temporal energies  $\mathcal{E}$  is obtained. In the frequency domain, the energies are decomposed into  $K$  frequency patterns using nonnegative matrix factorization (NMF). Each component of the NMF is transformed back into a collection signals, first by reconstructing temporal spectra using Wiener filtering, then by using inverse Fourier transforms. From the temporal collection of signals, a temporal network is obtained by using the inverse transformation  $\mathcal{T}^{-1}$  described in Section III-A.

- the energies  $\mathcal{E}$ , which read as  $e_{cf} = |s_{cf}|^2$
- the phases  $\Phi$ , which read as  $\phi_{cf} = \arg(s_{cf})$

The matrix  $\mathcal{E}$  is studied as a frequency-component map, exhibiting patterns in direct relation with the topology of the underlying graph. The phases of signals  $\Phi$  are used in the inverse Fourier transformation, when the collection of signals has to be retrieved from  $\mathcal{E}$ .

In the temporal case, the spectral analysis of signals is simply achieved by considering independently each time step. We denote by  $\mathcal{S} \in \mathbb{C}^{C \times F \times T}$  the spectral tensor, where  $\mathcal{S}^{(t)}$  corresponds to the spectra obtained at time  $t$ . The corresponding tensor  $\mathcal{E}$  represents the temporal energies.

As described in [10], the spectra are closely related to the network structures. In particular, the importance of high-energy components as well as low frequencies is highlighted, for instance for the structure in communities. Looking at the marginals of the temporal energies or magnitudes over frequencies and components is hence expected to give hints about the evolution of the structure of the temporal network over time. In the following, we will focus on the marginal of the energies over the components, denoted  $\mathcal{E}_c(f)$ , and over the frequencies, denoted  $\mathcal{E}_f(c)$ .

We also propose to use the spectral analysis to compare two network structures by computing the correlation between their spectra at each time step. The idea is to find among a set of parametric graph models the one that best fits the network at each time step. If we denote by  $\mathcal{S}_m$  the spectra obtained after transformation of an instance  $\mathcal{G}_m$  of a prescribed graph model, we can compute a correlation coefficient  $\rho^{(t)}$  between  $\mathcal{S}_m$  and  $\mathcal{S}^{(t)}$ . Generating several instances of the graph model gives us an average value of the correlation coefficient over several repetitions.

### C. Extraction of frequency patterns using nonnegative matrix factorization

1) *Nonnegative Matrix Factorization (NMF)*: Nonnegative matrix factorization (NMF) [26] is a linear regression technique, used to decompose a nonnegative matrix  $\mathbf{V}$  of dimension  $C \times T$ , i.e., a matrix whose terms are greater or equal to zero, into the product of two nonnegative matrices  $\mathbf{W}$  and

$\mathbf{H}$ . NMF leads to a reduction of the dimensionality of data, by extracting in the columns of  $\mathbf{W}$  patterns characterizing the data, and in the rows of  $\mathbf{H}$  the activation coefficients of each pattern along the time. The number of extracted patterns, or components, is denoted  $K$ . A common approach to achieve such a decomposition consists of solving the optimization problem:

$$(\mathbf{W}^*, \mathbf{H}^*) = \arg \min_{\mathbf{W}, \mathbf{H}} D(\mathbf{V} | \tilde{\mathbf{V}}) \quad (7)$$

with  $\tilde{\mathbf{V}} = \mathbf{W}\mathbf{H}$ , and where  $D$  is a dissimilarity measure between matrices given by  $D(\mathbf{V} | \mathbf{W}\mathbf{H}) = \sum_{c=1}^C \sum_{t=1}^T d(v_{ct} | \tilde{v}_{ct})$ , with  $d$  a dissimilarity measure between scalar values. Févotte et al. [33] proposed an algorithm to find a solution of the NMF where  $d$  is the  $\beta$ -divergence, a parametric function that encompasses the Euclidean distance, the generalized Kullback-Leibler divergence and the Itakura-Saito divergence as special cases.

Regularization of the activation coefficients can be added in order to smooth them, with the assumption that there is no abrupt variations in the structure from one time step to the next one. The optimization problem is then defined as the minimization of the fitting term, defined in Eq. (7), plus a term of temporal regularization:

$$P(\mathbf{H}) = \sum_{k=1}^K \sum_{t=2}^T d(h_{k(t-1)} | h_{kt}) \quad (8)$$

leading to:

$$(\mathbf{W}^*, \mathbf{H}^*) = \arg \min_{\mathbf{W}, \mathbf{H}} D(\mathbf{V} | \mathbf{W}\mathbf{H}) + \gamma P(\mathbf{H}) \quad (9)$$

where  $\gamma$  controls the regularization and is empirically fixed such that the activation coefficients highlight smoothness. In [34] and [35], smooth NMF has been introduced for  $\beta = 1$  and  $\beta = 0$ .

2) *NMF on spectra of graphs*: Several approaches have been proposed to adapt NMF to networks, either static [36] or temporal [19]. In the latter approach, the adjacency matrix is represented as a tensor and is decomposed using nonnegative tensor factorization (NTF) [37]. The drawback of this approach is that the adjacency matrix at each time step is represented

as the product of vectors, which is well-suited to highlight structure in communities but not adapted when the structure becomes more complex.

Following [30], we propose to use NMF to find patterns in spectra of the collections of signals, obtained from the transformation of the temporal network. By analogy with music analysis, where an audio sample is decomposed into several audio samples, separating for instance voice from the instrumental part [38], we would like to decompose the temporal network into temporal sub-networks, decomposing at each time step the global structure into several substructures. Furthermore, audio spectra share similarities with graph spectra, leading to use the Itakura-Saito divergence as measure of dissimilarity, given by

$$d_{IS}(v_{ct}|\tilde{v}_{ct}) = \frac{v_{ct}}{\tilde{v}_{ct}} - \log \frac{v_{ct}}{\tilde{v}_{ct}} - 1. \quad (10)$$

As the input in our case is the temporal spectra  $\mathcal{S}$ , represented as a tensor of dimension  $C \times F \times T$ , a small adaptation has to be performed before applying NMF. At each time instant  $t$ , the collection of spectra  $\mathcal{S}^{(t)}$  is represented as a vector  $v_t$  by successively adding end-to-end the columns of the matrix  $\mathcal{S}^{(t)}$ . For all  $t = 0, \dots, T-1$ , these vectors compose the columns of the matrix  $\mathbf{V}$ , of dimension  $(FC) \times T$ . The number of components  $K$  is set according to our expectations about the data, and the parameter  $\gamma$  is strictly positive to ensure smoothness in the activation coefficients.

3) *Identification of components*: NMF returns two matrices  $\mathbf{W}$  and  $\mathbf{H}$ : each column of  $\mathbf{W}$  represents the  $k$ th (normalized) frequency pattern, while the  $k$ th column of  $\mathbf{H}^T$ , gives the activation coefficients of the frequency pattern  $k$  at each time step. From  $w_k$ , a component-frequency map can be built by reshaping the vector into a matrix. To highlight how these structures are arranged in the temporal network, each component is transformed into temporal network: As described in [38], using NMF with the Itakura-Saito divergence provides means of reconstruction of the collection of signals corresponding to each component. As the matrix  $\mathbf{V}$  describes the amplitudes of signals, the frequency patterns obtained using NMF do not carry information about the phases, preventing the reconstruction into signals. By using Wiener filtering [38], the phases are reconstructed from the matrix  $\mathcal{S}^{(t)}$  directly, by decomposing each element according to the matrices  $\mathbf{W}$  and  $\mathbf{H}$ , enabling a reconstruction of signals corresponding to each component of the NMF. For each component  $k$ , the temporal spectrum  $\mathcal{S}^{(k)} \in \mathbb{C}^{C \times F \times T}$  is obtained such that its elements  $s_{cf}^{(k,t)}$  read as:

$$s_{cf}^{(k,t)} = \frac{w_{(cf)k} h_{kt}}{\sum_{l=1}^K w_{(cf)l} h_{lt}} s_{cf} \quad (11)$$

leading to a conservative decomposition of the tensor  $\mathcal{S}$ :

$$\mathcal{S} = \sum_{k=1}^K \mathcal{S}^{(k)} \quad (12)$$

The temporal spectrum of the component  $k$  is then a fraction of the original temporal spectrum. From  $\mathcal{S}^{(k)}$ , an inverse Fourier transformation is performed, leading to a collection of signals for each component  $k$  denoted by  $\mathcal{X}^{(k)} \in \mathbb{R}^{N \times N \times T}$ . Finally,

TABLE I  
GENERATION OF THE TOY TEMPORAL NETWORK: PROBABILITIES TO HAVE AN EDGE AT TIME  $t$

	$e \in E^p$	$e \notin E^p$
$e \in E_{t-1}$	0.99	0.8
$e \notin E_{t-1}$	0.2	0.01

the adjacency tensor  $\mathcal{A}^{(k)}$  describing the temporal network corresponding to the component  $k$  is obtained by using the inverse transformation  $\mathcal{T}^{-1}$  described in Section III.

Figure 1 summarizes the described framework developed for the study of temporal networks. In the next sections, we propose to illustrate it on a synthetic and a real-world temporal network. The implementation has been made using the Python language. The reader can refer to a dedicated webpage<sup>2</sup> to download the software implementation of the method.

#### IV. ILLUSTRATION ON A SYNTHETIC TEMPORAL NETWORK

##### A. Construction of the temporal network

The experiment is conducted first on a synthetic, yet realistic example, for which the ground truth is perfectly known. It consists of smooth transitions between different network structures, mixing communities, cycles and randomness.

Starting from  $N$  unconnected vertices, the algorithm adds or removes edges at each time instant according to a probability depending both on the presence or not of the edge at the previous time, and on a prescribed network structure. Let  $E_t$  be the set of edges at time  $t$ ,  $E^p$  the set of edges of a prescribed network structure, and  $e = (i, j)$  the edge between the vertices  $i$  and  $j$ . Table I gives the probability to have  $e$  in  $E_t$  according to the presence of  $e$  at the previous time (i.e.  $e \in E_{t-1}$ ) and the presence of  $e$  in the prescribed network structure ( $e \in E^p$ ). These probabilities are set up in order to generate a smooth transition between the initial structure and the prescribed network structure, and such that the structure at the end of the time interval is close to the one of the prescribed graph. Four networks are successively chosen as prescribed network structures, each structure being active for a time interval of 20 time steps:

- 0-20: Random linkage of vertices;
- 20-40: Network with 3 communities;
- 40-60: 4-ring lattice;
- 60-80: 4-ring lattice with 3 communities.

The prescribed structure is clearly visible thanks to a proper labeling of vertices: The blocks describe the communities, while the strong diagonal describes the regular lattice. These structures are nevertheless not completely defined, as noise due to the probabilities set in Table I remains present.

Figure 2 represents four snapshots of the temporal network at the end of the time intervals. For each snapshot, the adjacency matrix is displayed on the left and the spectra of graph-based signals on the right. These illustrations show that graph-based signatures can be attached to network structures thanks to the methodology outlined in Section II (for more details

<sup>2</sup><http://perso.ens-lyon.fr/ronan.hamon/index.php?page=software>

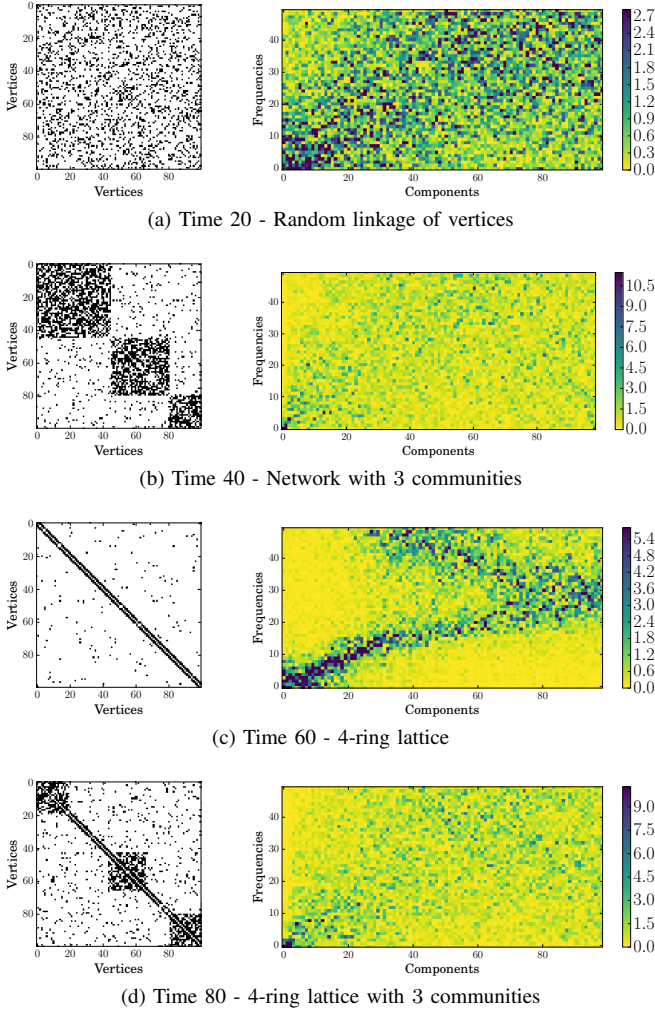


Fig. 2. Representation of the temporal network at the end of each time interval. (Left) Adjacency matrix. (Right) Component-frequency map obtained after spectral analysis. The color represents the intensity.

about such signatures, the reader is referred to [10]): random structures are represented by a frequency pattern whose energy is uniformly distributed, as in Figure 2a. The presence of specific structure in the network reveals particular patterns, such as high-energy low frequencies on the first components for the structure in communities (Figure 2b) and localized energies over specific frequencies for the regular structures (Figure 2c). Combination of both structures in the graph is preserved in the frequency pattern, as highlighted in Figure 2d.

### B. Network-based analysis

Network-based descriptors are used to analyze the structure of the network. Figure 3 plots three descriptors of the network at each time step: The number of edges (Figure 3a), the average clustering coefficient (Figure 3b) and the average length of shortest paths (Figure 3c) contribute to the identification of the four time intervals. The number of edges increases in the period 1 to form the random structure, characterized by a low average clustering coefficient and a low average length of shortest paths. Adding communities in period 2 increases the average clustering coefficient, as expected. On

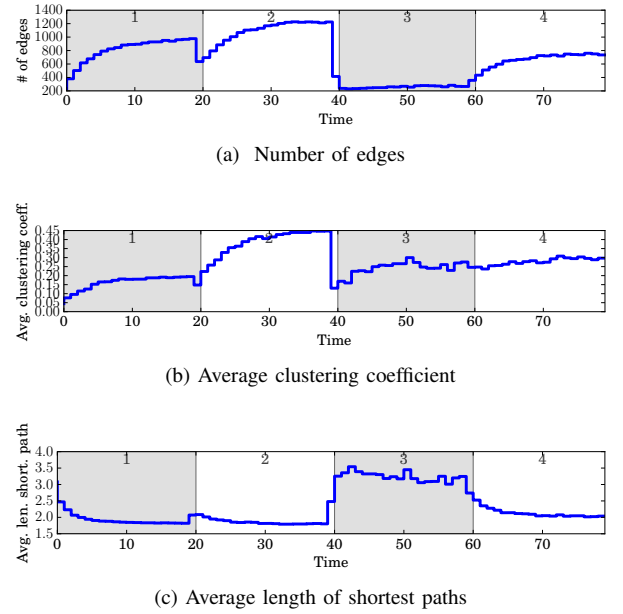


Fig. 3. Descriptors of the Toy Temporal Network over time. The alternating shaded regions correspond to the four different periods, whose label is given at the top of each region.

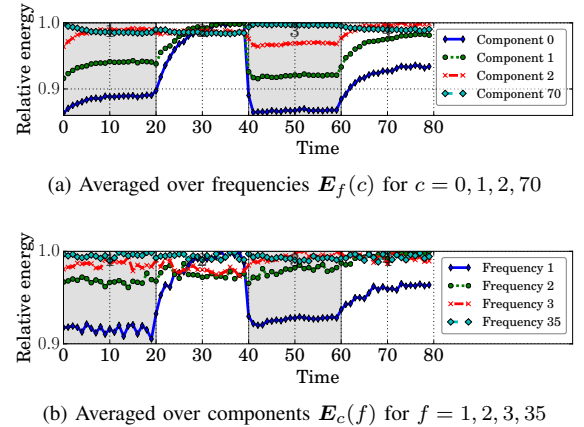


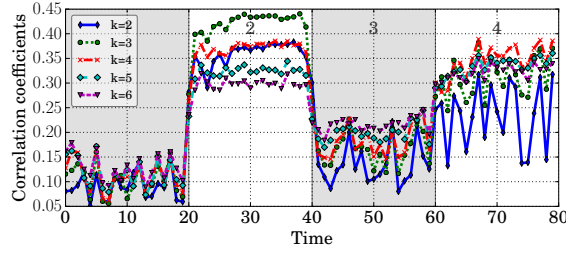
Fig. 4. Marginals of the energies of temporal spectra.

the contrary, the regular structure in period 3 decreases the clustering coefficient and increases the average length of shortest paths, but much less in proportion than the decrease of the number of edges would affect. As expected, adding communities in this regular structure, as it is done in period 4, significantly increases the average clustering coefficient and slightly decreases the average length of shortest paths. These network-based descriptors give good intuitions on the underlying structure of the temporal network, but turn out to be inefficient to exactly characterize the structure. Furthermore, mixture of different structures, as it appears in this model, are not explicitly revealed.

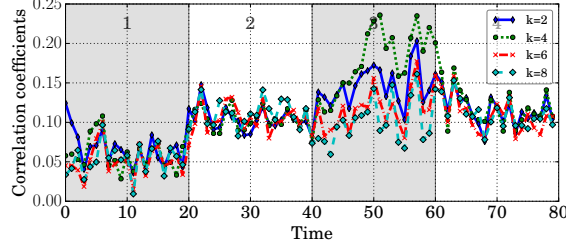
### C. Spectral analysis of graph-based signals

Figure 4 shows  $E_c(f)$  and  $E_f(c)$  for respectively  $f = 1, 2, 3, 35$  and  $c = 0, 1, 2, 70$ . These two figures reveal the predominance of the first components and low frequencies, to track an organization in communities of the network: the low





(a) Comparison with a network with  $k$  communities, for different values of  $k$ . The temporal network is highly correlated with a graph organized in communities during the period 2 and 4, as expected.



(b)  $k$ -regular lattice, for different value of  $k$ . The temporal network has the highest correlation with a 4-regular lattice during the period 3.

Fig. 5. Correlation between the temporal spectra at each time step and the spectra of networks with two specific network structures. The correlation is averaged over 20 repetitions.

frequencies for the first components have a greater energy than for other types of structure, as already remarked in Figure 2. The tracking of the other types of structures is nevertheless not visible using this representation.

In order to go further, the correlation between the temporal spectra of the toy temporal network and two network structures is studied. First, a structure in communities is observed, using a network model generating a random graph with a fixed number of communities. The comparison is done for a number of communities from 2 to 6, with 20 repetitions for each number of communities. Figure 5a shows the average correlation: the correlation is maximal during the periods 2 and 4, where the network is effectively structured in communities. During the period 2, the correlation is maximal when the network is structured in 3 communities, as expected. During the period 4, the number of communities is not clearly revealed, as the communities are not the only component of the network topology.

In Figure 5b, the temporal network is compared using the same method with a  $k$ -regular lattice, for  $k$  equals to 2, 4, 6 and 8. If the correlations are lower than previously, we can nevertheless notice that in period 3, the temporal network is correlated with a 4-regular lattice, which is the structure set in the prescribed network during this period.

The study of temporal spectra gives hence insights about the structure of the underlying temporal network, but this approach is limited by the lack of knowledge, in the general case, of the parametric graph models composing the temporal network.

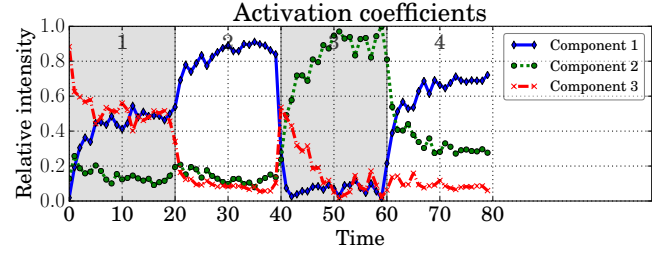
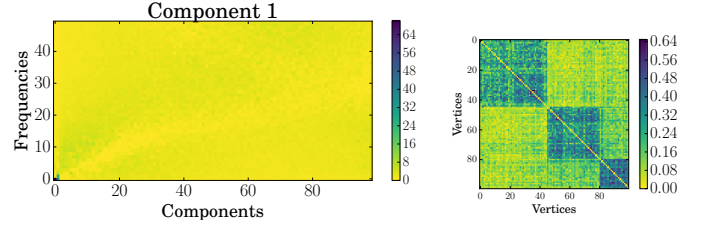
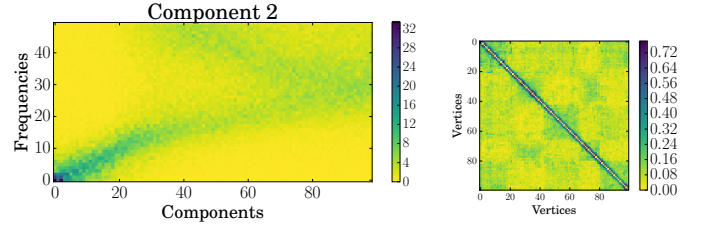


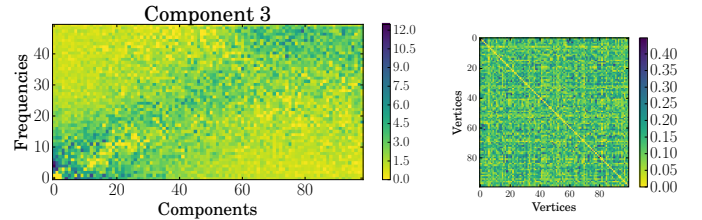
Fig. 6. Activation coefficients obtained after the nonnegative matrix factorization for the Toy Temporal Network, using  $K = 3$  and  $\gamma = 5$ . They correspond to the rows of the matrix  $H$ , normalized by the maximal value of  $H$ .



(a) Component 1



(b) Component 2



(c) Component 3

Fig. 7. (Left) Frequency patterns, obtained after reshaping the columns of  $W$  into matrices (arbitrary units). (Right) Adjacency tensor aggregated over time: for each component  $k$ ,  $A^{(k)} = \sum_{t=1}^T h_{kt} A^{(k,t)}$ .

#### D. Extraction of network structures

NMF is applied to the Toy Temporal Network defined in Section IV. The matrix  $V \in \mathbb{R}^{4950 \times 80}$  is decomposed using  $K = 3$  and  $\gamma = 5$ , which is the expected number of structures.

Figure 6 shows the activation coefficients of each component. We notice that the activation coefficients are consistent with the division of time introduced in the generation of the temporal network: All components have distinct levels of activation, corresponding to the four different structures used. The component 1 is active in periods 1, 2 and 4 with an almost constant level, the component 2 is mainly active in period 3, as well as in periods 1 and 4, and finally, the component 3 is active in period 1 and in period 2.

Jointly looking at the corresponding frequency patterns (and, more precisely, their connections) in Figure 7 (left plots)



and at the network structures observed in Figure 2, confirms the good match with the expected results: The component 1 looks like a structure in communities, the component 2 resembles a  $k$ -regular structure and the component 3 exhibits random structure. The structure in period 1 is then a mixture between a random structure and a structure in communities (as it happens, one single community), in period 2 only the structure in communities is present, in period 3 a regular structure described by component 2 and finally, the fourth period is a mixture between structure in communities and regular structure. Random structure is present in period 1 and in period 2 inside communities.

The right plots in Figure 7 display, for each component  $k$ , the aggregated adjacency matrix over time, obtained after the back transformation of the spectra into a temporal network. The sum is weighted by the activation coefficients given by the matrix  $H$ , in order to highlight visually the most significant patterns:

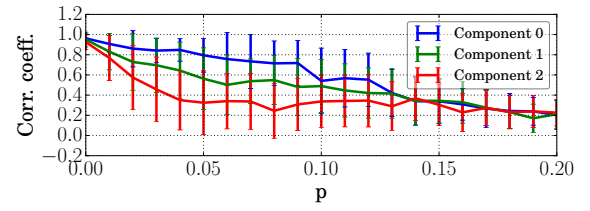
$$\mathbf{A}^{(k)} = \sum_{t=1}^T h_{kt} \mathbf{A}^{(k,t)} \quad (13)$$

This representation partly confirms the connections between spectra and structures as described above. We can notice that component 1 displays the three communities present in period 2, as well as communities corresponding to a regular structure in period 4. The actual communities in period 4 are caught in the component 2, which does not correspond to a regular structure, even if the diagonal, characterizing the  $k$ -regular lattice, is clearly dominant. Finally, the component 3 looks like a random matrix, as no structure is visible, at least through this representation.

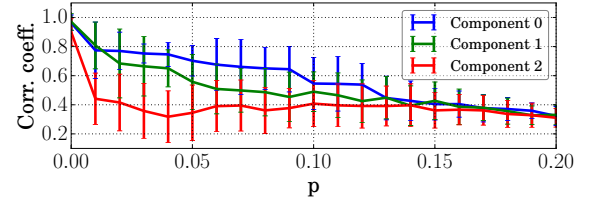
The decomposition of temporal networks using NMF enables to retrieve from the spectra the different structures composing the temporal network, and to detect when these structures are present, either alone or together.

#### E. Influence of the noise

A study of the robustness of the method in the presence of noise in the data is performed in the following. The presence of noise in the networks can be due to either the existence of a spurious link, or on the contrary to its absence. The procedure to assess the influence of the noise on the method of extraction of network structures is defined as follows: A uniform random noise is introduced in the data by flipping the entries of the adjacency matrix with a probability  $p$  at each time step. With such a process, edges are added or removed with the same probability. Then the noisy temporal network is transformed into signals, and components are extracted using the method described in Section III. The similarity between the components obtained in the noiseless situation and those obtained from the noisy temporal network is computed: The components are first paired such that each component refers to the same structure, in order to prevent any change in their order. Then the correlation coefficient between the vectors (corresponding to the rows of  $H$  and the columns of  $W$ ) is computed. This process is run for  $p = 0.00, 0.01, 0.02, \dots, 0.2$ , ranging from noiseless situation



(a) Activation coefficients



(b) Frequency patterns

Fig. 8. Study of the influence of the noise on the components obtained after extraction using the method described in Section III. The process described in Section IV-E is run for  $p = 0.00, 0.01, 0.02, \dots, 0.2$  and repeated 50 times.

to high level of noise. For each value of  $p$ , the operation is repeated 50 times.

Figure 8 displays the correlation coefficients for the activation coefficients (Figure 8a) and the frequency patterns (Figure 8b) of each component. Two main conclusions can be done: First, the correlation between noiseless and noisy components is high, meaning that the presence of noise does not affect the ability of the method to extract relevant network structures. Second, the correlation is slowly decreasing with respect to the value of  $p$ , which is in keeping with the expected behavior: If the temporal network is very noisy, the specific structures are less well-characterized as well as the extracted components. This is particularly true for the first two components, which represent respectively structure in communities and regular structures. As for the third component which represents randomness, the last observation is less obvious because the noise is also a source of randomness.

It is also worth noting that there are variations in the obtained components even in a noiseless situation ( $p = 0$ ). This is due to nonnegative matrix factorization, which provides an approximation of the solution, and is then subject to variability. These results nevertheless confirm that this variability exerts a low influence on the resulting component, as the correlation coefficients for both frequency patterns and activation coefficients are close to 1.

## V. A TEMPORAL NETWORK OF SOCIAL INTERACTIONS AS A REAL-WORLD EXAMPLE

### A. Description of the temporal network

The decomposition is applied to a real-world temporal network, describing social interactions between children in a primary school during two days in October 2009. During a 20-second interval, an edge exists between two individuals if a contact is recorded, measured by wearable RFID (Radio Frequency IDentification) sensors [39]. For our study, the data set is described by a temporal network, representing for each time step the aggregated contacts between individuals for

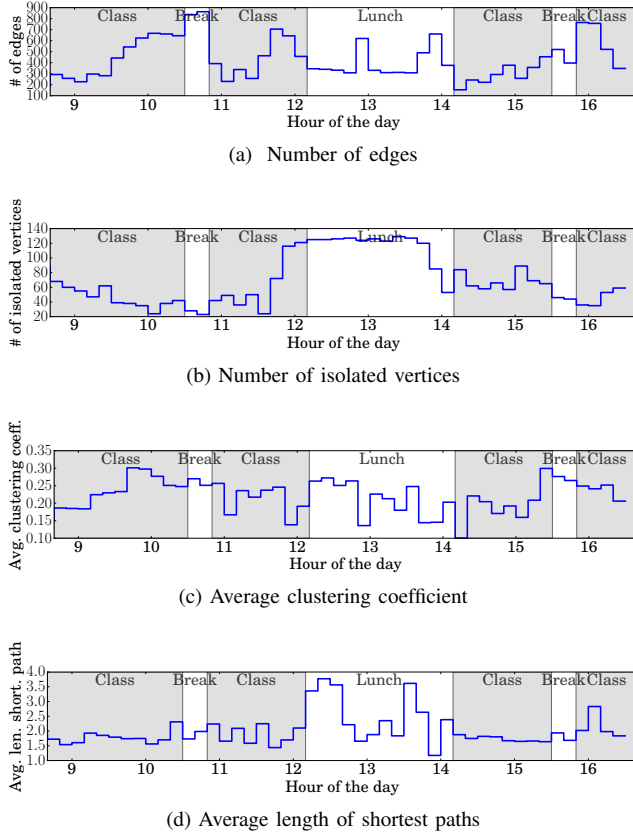


Fig. 9. Descriptors of the Primary School Temporal Network over time. The shaded regions correspond to class periods, while white regions correspond to breaks and lunch, according to the information given in [39]. No significant information is provided by the usual network-based descriptors

10 minutes. We restrained the analysis to the first day: 226 children and 10 teachers participated in the experiment, separated in five grades (from 1st grade to 5th grade), themselves separated in two classes.

Figure 9 shows the network-based descriptors classically used to characterize the temporal network. Figures 9a and 9b show that the number of edges in the temporal network, as well as the number of isolated vertices, is not constant over time, reflecting the real-world nature of data: during the lunch break, some children leave the school to have lunch at home. However, the network-based descriptors (Figures 9c and Figures 9d) do not provide significant insights about the structure of the temporal network.

### B. Spectral analysis of graph-based signals

Figure 10 shows the marginals of the energies of temporal spectra obtained from the transformation of the primary school temporal network. The energies of the low frequencies and of the first components are not equally distributed over time: this indicates changes in the global structure of the temporal network, that occur in break periods as well as during lunch. We can then divide the day into two main periods: the period where the children have class (shaded regions) and the breaks and lunch (white regions).

Comparing the Primary School Temporal Network with a network with communities using correlations between spectra

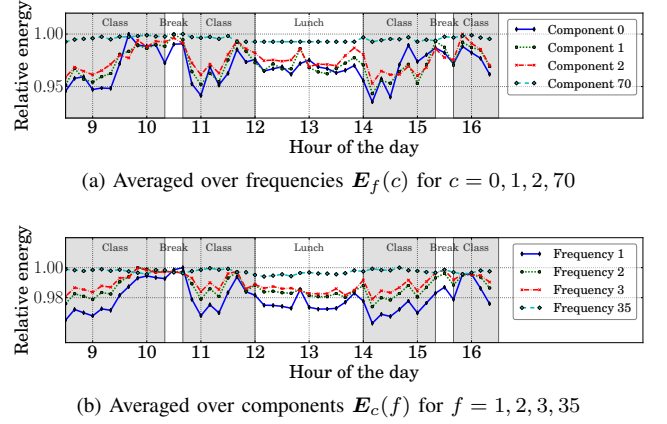


Fig. 10. Marginals of the energies of temporal spectra. The energies of the low frequencies and of the first components are not equally distributed over time, indicating changes in the global structure of the temporal network.

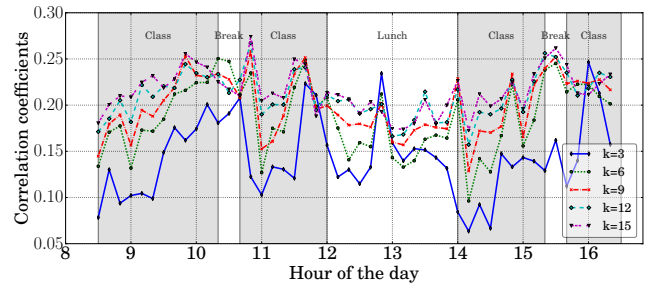


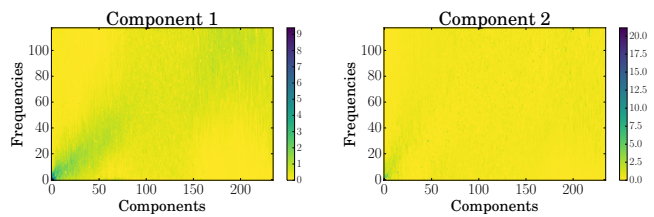
Fig. 11. Correlation between the temporal spectra at each time step and the spectra of a network with communities. Each line represent the number of communities, averaged over 20 repetitions. The temporal network is correlated with structure with a large number of communities during class periods, and with structure with a small number of communities during breaks and lunch.

shows that the former temporal network is correlated with the latter network involving a large number of communities (between 9 and 15) during class periods, a smaller one (between 3 and 6) during breaks and lunch periods (see Figure 11). This is consistent with the spatiotemporal trajectories given in [39], showing the location of the classes over time: during class periods, the classes are separated into different classrooms, while during breaks and lunch, the classes mix, yet in two distinct groups.

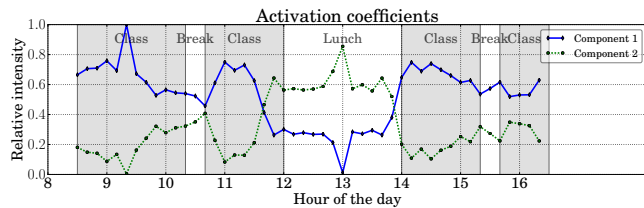
### C. Extraction of network structures

In the light of the above study of the temporal spectra, we can go further and decompose the Primary School Temporal Network into two temporal sub-networks. Furthermore, this allows for a quicker interpretation of the obtained components, as well as for an evidence of its ability to extract the significant components.

Figure 12 shows the results of the NMF on the Primary School Temporal Network, using  $K = 2$  and  $\gamma = 5$ . The school day is divided into three specific periods, according to the activation coefficients (Figure 12b). The first period occurs during class hours, where only the component 1 is mainly active. The second period groups together the breaks, during which the components 1 and 2 are significantly present. Finally, the period 3 concerns the lunch break, where the



(a) Frequency patterns, obtained after reshaping the columns of  $W$  into matrices (arbitrary units).



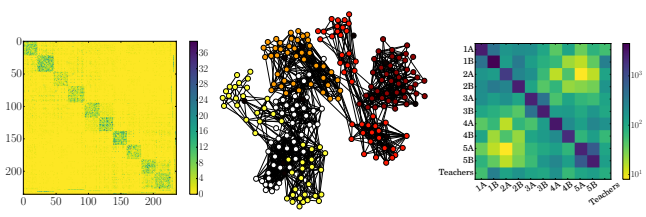
(b) Activation coefficients, corresponding to the rows of the matrix  $H$ , normalized by the maximal value of  $H$

Fig. 12. Results of the nonnegative matrix factorization for the Primary School Temporal Network, using  $K = 2$  and  $\gamma = 5$ . The activation coefficients are consistent with the schedule of the children in the primary school, as detailed in [39].

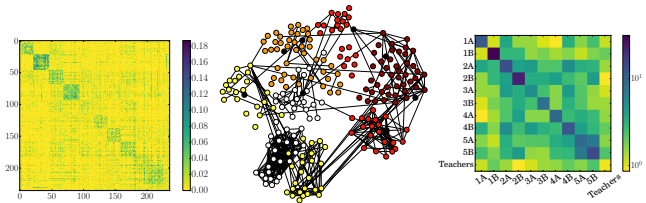
component 2 is dominant. This cutting is consistent with the ground truth, as described in [39]: only two or three classes have breaks at the same time, while the other ones stay in their respective classrooms. As for lunches, they are taken in two consecutive turns, preserving the structure in classes, with nevertheless a weaker intensity.

Figure 13 shows different representations of the temporal networks reconstructed from the components. The left figure shows the aggregated adjacency matrix over time, as described in Section IV-D. The vertices are ordered according to the classes of the children, from the youngest to the oldest. The middle figure shows the aggregated network, using the layout provided in [39], after thresholding of edges according to their weights. The color of dots indicates the grade of the children, while black dots represent the teachers. Finally, the right figure shows the contact matrix between classes, obtained by counting the number of edges inside and between the classes. A logarithmic scale is used to enhance the visualization. Figure 13a shows the original temporal network, while Figures 13b and 13c show respectively the component 1 and the component 2. We can easily observe that the component 1 describes the structure in classes, with higher density of edges inside classes than between classes. Conversely, the component 2 highlights a less structured network pattern, which looks like two communities, separating the youngest classes from the oldest. Those observations are consistent with the description of lunches mentioned above.

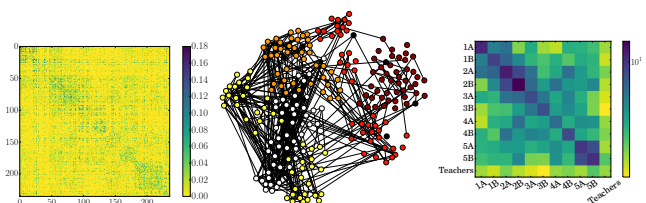
These results highlight the interest of decomposing a temporal network into several sub-networks, which can be studied independently of each another. Without prior knowledge, the different periods of activity in the primary school are displayed, and can then guide the analysis of the system by restricting the analysis over several time intervals.



(a) Original temporal network



(b) Component 1



(c) Component 2

Fig. 13. (Left) Aggregated adjacency matrix over time, weighted by the coefficients of  $H$ . (Middle) Network representation using the layout provided in [39], after thresholding of edges according to their weights. The color of dots indicates the grade of the children, while black dots represent the teachers. (Right) Grayscale-coded contact matrix between classes: each entry gives the number of contacts inside and between the classes. A logarithmic scale is used to enhance the visualization. The component 1 represents the structure in classes, while the component 2 describes the structure during the breaks and lunch.

## VI. CONCLUSION

We have proposed a novel method to track the structure of temporal networks over time using the duality between graphs and signals, as well as classical signal processing techniques, such as spectral analysis and nonnegative matrix factorization. At each time, the temporal network is represented by a graph which is transformed into a collection of signals. NMF is used to extract patterns in the energies of spectra of these signals; these patterns, whose activation coefficients vary over time, represent a specific structure of the underlying network. The effectiveness of the method has been demonstrated on a synthetic example, containing three types of structures as well as on a real-world network describing temporal contacts between children in a primary school.

These results provide insights in the characterization of temporal networks, but also call for further studies: In particular, it would be interesting to have a deeper look on how the reconstructed temporal networks are embedded in the original temporal network. Furthermore, it would be worth considering an iterative analysis, by studying the reconstructed temporal networks themselves, using the same process. Finally, one issue to the decomposition is how the number of components  $K$  should be chosen. The choice in this work has been guided by the intuition, according to the knowledge about the considered temporal networks. It could be nevertheless interesting









**Ronan Hamon** received the engineer degree from INSA Toulouse, France, in 2012, and the Ph.D. degree in Physics in 2015 at the Laboratoire de Physique de l'École Normale Supérieure de Lyon (ENS de Lyon), France. He is currently working as a Post-doctoral researcher in Laboratoire d'Informatique Fondamentale (LIF) in Marseille, France. His research interests are in graph signal processing, network theory, matrix factorization, and audio signal processing.



**Pierre Borgnat** (M'06), born in France in 1974, was received at the École Normale Supérieure de Lyon in 1994 and as a Professeur Agrégé in Physical Sciences in 1997. He received the Ph.D. degree in Physics and Signal Processing in 2002. He worked in 2003–2004 at the ISR, IST (Lisbon, Portugal). He is a CNRS research fellow since 2004 and works at the Physics Laboratory, ENS de Lyon. He is deputy director of IXXI (Complex System Institute of Rhône-Alpes) since 2015. His research interests

are in statistical signal processing, non-stationary signals, scaling phenomena and in processing of graph signals and complex networks. He works on several applications of these signal processing methods: Internet traffic modeling and measurements, fluid mechanics, analysis of social data, transportation studies.



**Patrick Flandrin** (M'85–SM'01–F'02) received the engineer degree from ICPI Lyon, France, in 1978, and the Doct.-Ing. and Docteur d'État degrees from INP Grenoble, France, in 1982 and 1987, respectively. He joined CNRS in 1982, where he is currently Research Director. Since 1991, he has been with the Signals, Systems and Physics Group, within the Physics Department at ENS de Lyon, France. He is currently President of GRETSI, the French Association for Signal and Image Processing. His research interests include mainly nonstationary signal processing (with emphasis on time-frequency and time-scale methods), scaling stochastic processes and complex systems. He published over 250 research papers and authored one monograph in those areas. Dr. Flandrin was awarded the Philip Morris Scientific Prize in Mathematics (1991), the SPIE Wavelet Pioneer Award (2001), the Prix Michel Monpetit from the French Academy of sciences (2001) and the Silver Medal from CNRS (2010). Past Distinguished Lecturer of the IEEE Signal Processing Society (2010–2011), he is a Fellow of the IEEE (2002) and of EURASIP (2009), and he has been elected member of the French Academy of sciences in 2010.

signal processing (with emphasis on time-frequency and time-scale methods), scaling stochastic processes and complex systems. He published over 250 research papers and authored one monograph in those areas. Dr. Flandrin was awarded the Philip Morris Scientific Prize in Mathematics (1991), the SPIE Wavelet Pioneer Award (2001), the Prix Michel Monpetit from the French Academy of sciences (2001) and the Silver Medal from CNRS (2010). Past Distinguished Lecturer of the IEEE Signal Processing Society (2010–2011), he is a Fellow of the IEEE (2002) and of EURASIP (2009), and he has been elected member of the French Academy of sciences in 2010.



**Céline Robardet** received the Ph.D. degree in Computer Science in 2002 from University Lyon 1. Then, she joined the French national institut of applied sciences of Lyon as associate professor and became full professor in 2014. She is the leader of the DM2L (Data Mining and Machine Learning) research group in LIRIS lab. Her main research interests include several aspects of data mining and combinatorial optimization problems, such as clustering analysis, pattern extraction under constraints and complex dynamic network analysis.